Gene 507 (2012) 112–118

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

New features of Asian *Crassostrea* oyster mitochondrial genomes: A novel alloacceptor tRNA gene recruitment and two novel ORFs

Xiangyun Wu [a], Xiaoling Li [a,b], Lu Li [a], Xiaodong Xu [a,b], Jianjun Xia [a], Ziniu Yu [a,*]

[a] Key Laboratory of Marine Bio-resource Sustainable Utilization, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China
[b] Graduate School of the Chinese Academy of Sciences, 19A Yuquan Road, Beijing, 100049, China

ARTICLE INFO

Article history:
Accepted 18 July 2012
Available online 27 July 2012

Keywords:
Crassostrea
Mitochondrial genome
tRNA gene recruitment
Oyster
atp8

ABSTRACT

A feasible way to perform evolutionary analyses is to compare characters divergent enough to observe significant differences, but sufficiently similar to exclude saturation of the differences that occurred. Thus, comparisons of invertebrate mitochondrial (mt) genomes at low taxonomic levels can be extremely helpful in investigating patterns of variation and evolutionary dynamics of genomes, as intermediate stages of the process may be identified. Fortunately, in this study, we newly sequenced the mt genome of the eighth member of Asian *Crassostrea* oysters which can provide necessary intermediate characters for us to believe that the variation of *Crassostrea* mt genomes is considerably greater than previously acknowledged. Several new features of Asian *Crassostrea* oyster mitochondrial genomes were revealed, and our results are particularly significant as they 1) suggest a novel model of alloacceptor tRNA gene recruitment, namely "vertical" tRNA gene recruitment, which can be successfully used to explain the origination of the unusually additional *trnK* and *trnQ* genes (annotated as $trnK_2$ and $trnQ_2$ respectively) in the mt genomes of the five Asian oysters, and we speculate that this recruitment progress may be a common phenomenon in the evolution of the tRNA multigene family; 2) reveal the existence of two additional, lineage-specific, mtDNA-encoded genes that may originate from duplication of *nad2* followed by rapid evolutionary change. Each of these two genes encodes a unique amino terminal signal peptide, thus each might possess an unknown function; and 3) identify for the first time the *atp8* gene in oysters. The present study thus gives further credence to the comparison of congeneric bivalves as a meaningful strategy to investigate mt genomic evolutionary trends in genome organization, tRNA multigene family, and gene loss and/or duplication that are difficult to undertake at higher taxonomic levels. In particular, our study provides new evidence for the identification and characterization of ORFs in the "non-coding region" of animal mt genomes.

1. Introduction

Bivalve mitochondrial (mt) genomes are highly variable in size, and gene organization is known to vary extensively even among species from the same family, *e.g.*, Pectinidae (Ren et al., 2010b; Wu et al., 2009), Unionidae (Breton et al., 2009) and Veneridae (Xu et al., 2010). The striking differences in mt genome organization of bivalves may reflect their potentially weak ability to maintain a stable genome. This view is consistent with recent conclusions that mt genomes of many animal phyla may be considerably more plastic than previously thought, and that the conserved genome organization of vertebrates reflects a derived stabilization of the mt genome rather than an ancestral feature (Gissi et al., 2008). Therefore, comparisons of invertebrate mt genomes at low taxonomic levels can be extremely helpful in investigating patterns of variation and evolutionary dynamics of mt genomes. For example, the investigation of gene rearrangements in closely related species can be very useful for the discovery of the mechanisms underlying genome evolution, as intermediate stages of the process may be identified (Kurabayashi et al., 2008; Xu et al., 2012). The existence of two or multiple copies of a tRNA gene is commonly seen in mt genomes. For example, there are two distinct *trnM* genes in nearly all available bivalve mt genomes, two *trnV* genes in the marine clam *Venerupis philippinarum*, and two *trnD* genes in the scallop *Mizuhopecten yessoensis* (Wu et al., 2009) and in four cephalopods (*Dosidicus gigas*, *Sthenoteuthis oualaniensis*, *Todarodes pacificus*, *Watasenia scintillans*) (Akasaki et al., 2006; Staaf et al., 2010; Yokobori et al., 2004). We consider the appearance of two tRNA genes that sporadically occur in different animal lineages as occasional events, and the origin of the additional tRNA genes is relatively unstudied. In most cases, the two tRNA gene copies show high divergences in their primary sequences, suggesting that they were not created by a recent gene duplication event. However, the signature of specific processes

Abbreviations: *atp6* and *atp8*, ATPase subunit 6 and 8 genes; *cob*, cytochrome *b* gene; *cox1–3*, cytochrome *c* oxidase subunits I–III genes; *nad1–6* and *4L*, NADH dehydrogenase subunits 1–6 and 4L genes; rRNA, ribosomal RNA; *rrnL* and *rrnS*, large and small subunits of ribosomal RNA genes; tRNA, transfer RNA; *trnM*, methionine transfer RNA gene..

* Corresponding author. Tel./fax: +86 20 8910 2507.
E-mail address: carlzyu@scsio.ac.cn (Z. Yu).

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2012.07.032

such as tRNA gene recruitment can be most easily recognized if such events have occurred recently and thus will be more easily seen in conspecific and/or congeneric mt genomes (Gissi et al., 2008; Lavrov and Lang, 2005).

Crassostrea oysters are distributed worldwide, and nearly 20 species in this genus are recognized (Huber, 2010), seven of them are reported in China (Xia, 2008; Xia and Yu, 2009; Zhuang, 2001), namely *Crassostrea angulata* (Lamarck, 1819), *Crassostrea ariakensis* (Fujita, 1913), *Crassostrea gigas* (Thunberg, 1793), *Crassostrea hongkongensis* Lam & Morton, 2003, *Crassostrea nippona* (Seki, 1934), *Crassostrea sikamea* (Amemiya, 1928), and *Crassostrea iredalei* (Faustino, 1932). Our previous study indicated that mt genomes of seven *Crassostrea* species (six from Asia and one from America) exhibit several unusual features compared to those of other bivalves (Wu et al., 2010), including *rrnS* duplication in six Asian oysters and *rrnL* splitting in both Asian and American members, rearrangements of tRNA genes between Asian and American oysters, and the loss of the *atp8* gene in all oysters. Ren et al. (2010a) performed a similar comparison and concluded that the complete conservation of gene order in the six Asian *Crassostrea* species is highly unusual. However, this viewpoint concerning the conserved gene arrangement in Asian oyster mt genomes was soon broken by a gene rearrangement event in *C. nippona* where the "*trnG* + MNR" cluster is moved upstream of the gene cluster "*trnN–rrnS*$_2$*–trnY–atp6*" (Yu and Li, 2012).

Meanwhile, the available *C. nippona* mt genome indicates the end of a task to sequence the complete mtDNA of all seven described *Crassostrea* oysters in China. These genomic sequences did successfully help us in understanding the relationship of this commercially important oyster group (Ren et al., 2010a; Wu et al., 2010), but provided little information concerning their genomic evolution. Evolutionary patterns of several characters only found within *Crassostrea* mt genomes are still unknown, such as the loss of the *atp8* gene, the presence of the additional *trnK* and *trnQ* genes in four oysters but the absence of them in others, and the unusual elongated genome size of *C. iredalei* (Wu et al., 2010). It's well known that a feasible way to perform evolutionary analyses is to compare characters divergent enough to observe significant differences, but sufficiently similar to exclude saturation of the differences that occurred. Fortunately, in this study, we will report a new mt genome which can provide necessary intermediate characters for us to trace the origin and evolutionary pathway of the tRNA gene family and novel ORFs with unknown functions in Asian *Crassostrea* mt genomes. This species, named *Crassostrea dianbaiensis*, is the eighth member of *Crassostrea* oysters in China. Briefly speaking, *C. dianbaiensis* is morphologically similar and phylogenetically close to *C. iredalei*, two oysters with sympatric distributions. These two species can be differentiated by their stable distinct adductor scar and obvious genetic divergence revealed by partial *cox1* and *rrnL* sequences, as well as the evidence from the entire mtDNA sequence shown in this study. In order to obey the basic rule of taxonomic study, in this study, we temporarily named this unpublished species as *Crassostrea* sp.DB, here "DB" is the abbreviation of Dianbai County (Guangdong Province, China) which is the location where the specimens are collected from.

2. Materials and methods

2.1. Specimens, DNA extraction, PCR amplification and sequencing

Whole genomic DNA was extracted from the adductor muscles of *Crassostrea* sp.DB using a TIANamp Marine Animal DNA kit (Tiangen, Beijing). Short fragments from genes *cox1*, *cob*, *atp6*, and *nad5* were amplified by PCR with universal primer pairs (Wu et al., 2010) designed based on the alignment of the published oyster mt genome sequences. Based on the sequences of these fragments, long-PCR primers were designed and employed to amplify overlapping segments of each entire mt genome (Supplementary Table 1, Supplementary Fig. 1). PCR reactions were performed in a 25 μL volume with 0.5 μL of template DNA (approximately 30 ng), 2.5 μL of 10× LA-buffer (Mg^{2+} plus), 0.5 μL of 10 mM dNTP mix, 1 μL of each primer (10 μM), and 0.25 μL (1 U) of LA*Taq* polymerase (Takara, Dalian, China). The PCR reactions were performed on an ABI Veriti thermal cycler (Applied Biosystems, California, USA) with the following parameters: pre-denaturation at 94 °C for 1 min followed by 35 cycles of 94 °C for 20 s, 45–55 °C annealing temperature for 20 s, extension at 68 °C for 2–8 min, and a final extension step at 72 °C for 10 min. PCR products were separated by electrophoresis on a 1.0% agarose gel, purified with a QIAquick PCR Purification kit (QIAGEN, California, USA) and bi-directionally sequenced using a primer-walking strategy on an ABI 3730xl DNA Sequencer (Applied Biosystems, California, USA).

2.2. Sequence assembly, annotation and analysis

Sequences were assembled using a SeqMan program (DNAstar, Madison, Wisconsin). Manual examinations were applied to ensure correct assembly. Protein coding genes (PCGs) and rRNA genes were identified by comparison with homologous mt DNA sequences of previously published oysters. Most tRNA genes were identified by the program tRNAscan-SE1.21 (Lowe and Eddy, 1997), with others identified by comparison with homologous tRNA gene sequences of previously published oysters. Mt genome maps were generated by CGview (Stothard and Wishart, 2005), and a complete mt genome has been deposited in GenBank under accession number JQ060958.

MEGA 4 (Tamura et al., 2007) was used for sequence alignments and to calculate the proportion of both nucleotide and amino acid differences (*p*-distances). Sequence similarity searches were performed in GenBank using blastn. Examination of ORFs was performed using the NCBI ORF Finder program (http://www.ncbi.nlm.nih.gov/projects/gorf/). The presence and location of a transmembrane hydrophobic helix in *atp8* protein were investigated using the TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). The protein domain was predicted with the Simple Modular Architecture Research Tool (SMART) program (http://smart.emblheidelberg.de/). In order to infer the type of evolutionary pressure acting on the two newly identified ORFs, we sequenced the fragments containing the two ORFs from five additional individuals of *C. iredalei* and *Crassostrea* sp.DB, separately. Values of non-synonymous changes per non-synonymous site (dN) and that of synonymous changes per synonymous site (dS) were calculated by MEGA 5.0 using the Jukes–Cantor model.

2.3. Phylogenetic reconstruction

Phylogenetic relationships of eight Asian *Crassostrea* oysters were reconstructed based on both nucleotide and amino acid sequences of 13 PCGs, using the American oyster *Crassostrea virginica* as the outgroup. Each gene was translated into an amino acid sequence using the invertebrate mt genetic code in MEGA, and aligned based on its amino acid sequence using default settings. The alignment was back-translated to the corresponding nucleotide sequences, and the final nucleotide sequences of each gene were then concatenated into single contigs for phylogenetic analyses. Maximum parsimony (MP) and maximum likelihood (ML) were employed for phylogenetic reconstructions. MP analyses were performed using PAUP* 4.0b10 (Swofford, 2002), with a total of 1000 random addition searches using TBR. Bootstrap support (BP) values were calculated from 1000 bootstrap replicates with 10 random additions per replicate in PAUP. ML analyses were performed using both PhyML 3.0 online execution (http://www. atgc-montpellier.fr/phyml) and MEGA 4. The GTR model with its parameter for the concatenated nucleotide dataset was determined by Modeltest (Posada and Crandall, 1998) based on the Akaike information criterion. Bootstrap support for ML trees was calculated using 500 bootstrap replicates. Phylogenetic relationships among *nad2* genes in all available *Crassostrea* mt genomes and the two ORFs found in *C. iredalei* and *Crassostrea* sp.DB were inferred